

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

特開平11-133992

(43) 公開日 平成11年(1999) 5月21日

(51) Int. Cl. °	識別記号	F 1		
G10L 3/00	535	G10L 3/00	535	
	541		541	
G06T 7/00		3/02	301	Z
G10L 3/02	301	9/10	301	C
9/10	301	G06F 15/70	460	B
審査請求 未請求 請求項の数24 O L (全14頁)				

(21) 出願番号 特願平9-300979

(22) 出願日 平成9年(1997)10月31日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 岩橋 直人

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 包 洪長

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 本田 等

東京都品川区北品川6丁目7番35号 ソニー株式会社内

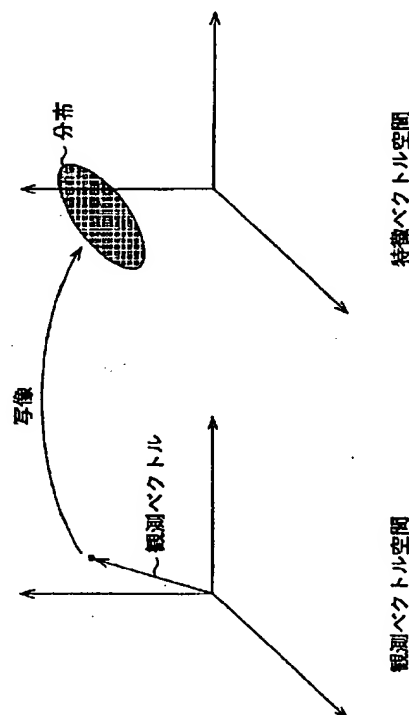
(74) 代理人 弁理士 稲本 義雄

(54) 【発明の名称】 特徴抽出装置および特徴抽出方法、並びにパターン認識装置およびパターン認識方法

(57) 【要約】

【課題】 音声認識や画像認識における認識率を向上させる。

【解決手段】 入力データとしての観測ベクトル a は、観測ベクトル空間上のある点を表すが、その観測ベクトル a が、特徴ベクトル空間に、その特徴ベクトル空間における広がりをもつ分布として写像され、その分布を表す特徴分布パラメータが求められる。そして、その特徴分布パラメータに基づいて、入力データがパターン認識される。



【特許請求の範囲】

【請求項 1】 入力データの特徴量を抽出する特徴抽出装置であって、

前記入力データの特徴量の空間に、前記入力データを写像したときに得られる分布を表す特徴分布パラメータを算出する算出手段を備えることを特徴とする特徴抽出装置。

【請求項 2】 前記算出手段は、正規確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 1 に記載の特徴抽出装置。

【請求項 3】 前記算出手段は、対数正規確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 1 に記載の特徴抽出装置。

【請求項 4】 前記算出手段は、離散確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 1 に記載の特徴抽出装置。

【請求項 5】 前記算出手段は、ファジー分布を表す特徴分布パラメータを算出することを特徴とする請求項 1 に記載の特徴抽出装置。

【請求項 6】 前記算出手段は、前記入力データの複数種類の特徴量のうちの、少なくとも 1 種類の特徴量の空間について、前記特徴分布パラメータを算出することを特徴とする請求項 1 に記載の特徴抽出装置。

【請求項 7】 入力データの特徴量を抽出する特徴抽出方法であって、

前記入力データの特徴量の空間に、前記入力データを写像したときに得られる分布を表す特徴分布パラメータを算出することを特徴とする特徴抽出方法。

【請求項 8】 入力されたデータのパターンを、所定数のクラスのうちのいずれかに分類することにより認識するパターン認識装置であって、

前記入力データの特徴量の空間に、前記入力データを写像したときに得られる分布を表す特徴分布パラメータを算出する算出手段と、

前記特徴分布パラメータを、前記所定数のクラスのうちのいずれかに分類する分類手段とを備えることを特徴とするパターン認識装置。

【請求項 9】 前記算出手段は、正規確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 10】 前記算出手段は、対数正規確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 11】 前記算出手段は、離散確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 12】 前記算出手段は、ファジー分布を表す特徴分布パラメータを算出することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 13】 前記算出手段は、前記入力データの複

数種類の特徴量のうちの、少なくとも 1 種類の特徴量の空間について、前記特徴分布パラメータを算出し、前記分類手段は、残りの種類の特徴量、および前記特徴分布パラメータを、前記所定数のクラスのうちのいずれかに分類することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 14】 前記分類手段は、正規確率分布関数を用いて、前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 15】 前記分類手段は、多項式関数を用いて、前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 16】 前記分類手段は、HMM (Hidden Markov Model) 法を用いて、前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 17】 前記分類手段は、ニューラルネットワークを用いて前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 18】 前記入力データは、音声データであることを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 19】 前記算出手段は、ノイズを用いて、前記特徴分布パラメータを算出することを特徴とする請求項 8 に記載のパターン認識装置。

【請求項 20】 前記算出手段は、正規確率分布を表す特徴分布パラメータを算出することを特徴とする請求項 18 に記載のパターン認識装置。

【請求項 21】 前記分類手段は、HMM (Hidden Markov Model) 法を用いて、前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 18 に記載のパターン認識装置。

【請求項 22】 前記分類手段は、ニューラルネットワークを用いて前記特徴分布パラメータが、前記所定数のクラスのうちのいずれに属するかを判定することを特徴とする請求項 18 に記載のパターン認識装置。

【請求項 23】 前記算出手段は、前記音声データのパワースペクトラム空間またはケプストラム空間における分布を表す特徴分布パラメータを算出することを特徴とする請求項 18 に記載のパターン認識装置。

【請求項 24】 入力されたデータのパターンを、所定数のクラスのうちのいずれかに分類することにより認識するパターン認識方法であって、

前記入力データの特徴量の空間に、前記入力データを写像したときに得られる分布を表す特徴分布パラメータを算出し、

前記特徴分布パラメータを、前記所定数のクラスのうちのいずれかに分類することを特徴とするパターン認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、特徴抽出装置および特徴抽出方法、並びにパターン認識装置およびパターン認識方法に関し、例えば、特に、ノイズ環境下において音声認識を行う場合などに用いて好適な特徴抽出装置および特徴抽出方法、並びにパターン認識装置およびパターン認識方法に関する。

【0002】

【従来の技術】図9は、従来のパターン認識装置の一例の構成を示している。

【0003】特徴抽出部101には、パターン認識対象としての観測ベクトルが入力される。特徴抽出部101では、観測ベクトルから、その特徴量を表す特徴ベクトルが求められ、識別部102に供給される。識別部102では、特徴抽出部101からの特徴ベクトルに基づいて、入力された観測ベクトルが、所定数のクラスのうちのいずれに属するかが識別される。

【0004】即ち、図9のパターン認識装置が、例えば、音声認識装置を構成している場合には、観測ベクトルとして、所定の時間単位（以下、適宜、フレームという）ごとの音声データが、特徴抽出部101に入力される。特徴抽出部101では、観測ベクトルとしての音声データが音響分析され、これにより、パワースペクトラムや、ケプストラム係数、線形予測係数、その他の音声の特徴量としての特徴ベクトルが抽出される。この特徴ベクトルは、識別部102に供給され、そこでは、特徴ベクトルが、所定数のクラスのうちのいずれかに分類され、その分類結果が、観測ベクトルとしての音声データの認識結果として出力される。

【0005】ここで、識別部102において、特徴ベクトルが、所定数のクラスのうちのいずれに属するかを識別する方法としては、例えば、マハラノビス識別関数や、混合正規分布関数、多項式関数を用いる方法、HMM法を用いる方法、ニューラルネットワークを用いる方法などが知られている。

【0006】なお、上述のような音声認識の手法については、例えば、L.Rabiner, B-H. Juang共著、古井監訳、「音声認識の基礎（上）（下）」、NTTアドバンステクノロジー（株）、1995などに、その詳細が開示されている。また、パターン認識一般に関しては、例えば、R.Duda and P.Hart, "Pattern Classification and Scene Analysis", John Wiley & Sons 1973などに、その詳細が開示されている。

【0007】ところで、パターン認識を行う場合に、そのパターン認識対象である観測ベクトル（入力パターン）には、一般に、ノイズが含まれる。即ち、例えば、

音声認識を行う場合に入力される観測ベクトルとしての音声には、ユーザが発話を行う環境下におけるノイズ

（例えば、他人の話し声や、自動車の騒音など）が含まれる。また、例えば、画像認識を行う場合に入力される観測ベクトルとしての画像には、その画像の撮影時の環境下におけるノイズ（例えば、霧や雨などの気象条件によるノイズ、画像を撮影するカメラのレンズの収差によるノイズなど）が含まれる。

【0008】そこで、例えば、ノイズ環境下において音声認識する場合の特徴量（特徴ベクトル）の抽出方法の1つに、スペクトルサブトラクション（Spectral Subtraction）と呼ばれるものがある。

【0009】スペクトルサブトラクションでは、音声の発話が行われる前の入力（音声区間の前の入力）を、ノイズとして、そのノイズの平均スペクトルが算出される。そして、音声の入力があると、その音声から、ノイズの平均スペクトルが差し引かれ（Subtract）、その残りを、真の音声成分として、特徴ベクトルが算出される。

【0010】なお、スペクトルサブトラクションについては、例えば、S.F.Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No. 2 1979や、P.Lockwood and J. Boudy, "Experiments with a Nonlinear Spectral Subtractor, Hidden Markov Models and the projection, for robust speech recognition in cars", Speech Communication, Vol. 11, 1992などに、その詳細が開示されている。

【0011】

【発明が解決しようとする課題】ところで、図9のパターン認識装置における特徴抽出部101では、図10に示すように、観測ベクトル空間上の、ある点を表す観測ベクトル a を、特徴ベクトル空間上に写像し、これにより、その特徴ベクトル空間上の、対応する点を表す特徴ベクトル y に変換する処理が行われると考えることができる。

【0012】従って、特徴ベクトル y は、特徴ベクトル空間上の、ある1点（観測ベクトル a に対応する点）を表す。なお、図10では、観測ベクトル空間および特徴ベクトル空間のいずれも3次元としてある。

【0013】スペクトルサブトラクションでは、観測ベクトル a から、ノイズの平均スペクトル成分が取り除かれて、特徴ベクトル y が算出されるが、この特徴ベクトル y は、上述したように、特徴ベクトル空間上の1点であるため、ノイズの平均的な性質を考慮したものとはなっているが、ノイズの分散などの不規則な性質を考慮したものとはなっていない。

【0014】このため、特徴ベクトル y は、観測ベクトル a の特徴を十分に表現しているとはいえず、そのような特徴ベクトル y では、高い認識率を得るのが困難であ

った。

【0015】本発明は、このような状況に鑑みてなされたものであり、認識率を向上させることができるようにするものである。

【0016】

【課題を解決するための手段】請求項1に記載の特徴抽出装置は、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出する算出手段を備えることを特徴とする。

【0017】請求項7に記載の特徴抽出方法は、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出することを特徴とする。

【0018】請求項8に記載のパターン認識装置は、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出する算出手段と、特徴分布パラメータを、所定数のクラスのうちのいずれかに分類する分類手段とを備えることを特徴とする。

【0019】請求項24に記載のパターン認識方法は、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出し、特徴分布パラメータを、所定数のクラスのうちのいずれかに分類することを特徴とする。

【0020】請求項1に記載の特徴抽出装置において、算出手段が、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出するようになされている。

【0021】請求項7に記載の特徴抽出方法において、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出するようになされている。

【0022】請求項8に記載のパターン認識装置においては、算出手段は、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出し、分類手段は、特徴分布パラメータを、所定数のクラスのうちのいずれかに分類するようになされている。

【0023】請求項24に記載のパターン認識方法においては、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータを算出し、特徴分布パラメータを、所定数のクラスのうちのいずれかに分類するようになされている。

【0024】

【発明の実施の形態】図1は、本発明を適用した音声認識装置の一実施の形態の構成例を示している。

【0025】フレーム化部1には、認識対象のデジタル音声データが入力されるようになされている。フレーム化部1は、例えば、図2に示すように、そこに入力される音声データを、所定の時間間隔（例えば、10ms

など）で取り出し（フレーム化し）、その取り出した音声データを、1フレームのデータとして出力するようになされている。フレーム化部1が出力する1フレーム単位の音声データは、そのフレームを構成する時系列の音声データそれぞれをコンポーネントとする観測ベクトル a として、特徴抽出部2に供給されるようになされている。

【0026】ここで、以下、適宜、第 t フレームの音声データである観測ベクトルを、 $a(t)$ と表す。

【0027】特徴抽出部2（算出手段）は、フレーム化部1からの観測ベクトル a としての音声データを音響分析し、これにより、その特徴量を抽出するようになされている。即ち、特徴抽出部2は、例えば、観測ベクトル a としての音声データをフーリエ変換し、そのパワースペクトラムを求め、そのパワースペクトラムの各周波数成分をコンポーネントとする特徴ベクトル y を算出する。なお、パワースペクトラムの算出方法は、フーリエ変換によるものに限定されるものではない。即ち、パワースペクトラムは、その他、例えば、いわゆるフィルタバンク法などによって求めることも可能である。

【0028】さらに、特徴抽出部2は、観測ベクトル a としての音声データに含まれる真の音声を、その特徴量の空間（特徴ベクトル空間）に写像したときに得られる、その特徴ベクトル空間上の分布を表すパラメータ（以下、特徴分布パラメータという） Z を、算出した特徴ベクトル y に基づいて算出し、識別部3に供給するようになされている。

【0029】即ち、特徴抽出部2では、図3に示すように、観測ベクトル空間上の、ある点を表す観測ベクトル a を、特徴ベクトル空間に、その特徴ベクトル空間における広がりを持する分布として写像したときの、この分布を表すパラメータが、特徴分布パラメータとして算出されて出力される。

【0030】なお、図3では、観測ベクトル空間および特徴ベクトル空間のいずれも3次元としてあるが、観測ベクトル空間および特徴ベクトル空間の次元数は、3次元に限定されるものではなく、さらに、それらの次元数は同一である必要もない。

【0031】識別部3（分類手段）は、特徴抽出部2からの特徴分布パラメータ（系列）を、所定数のクラスのうちのいずれかに分類し、その分類結果を、入力された音声の認識結果として出力するようになされている。即ち、識別部3は、例えば、所定数 K の単語それぞれに対応するクラスかどうかを識別するための識別関数を記憶しており、各クラスの識別関数の値を、特徴抽出部2からの特徴分布パラメータを引数として計算する。そして、その関数値の最も大きいクラス（ここでは、単語）が、観測ベクトル a としての音声の認識結果として出力される。

【0032】次に、その動作について説明する。

【0033】フレーム化部1には、認識対象のデジタル音声データが入力され、そこでは、音声データがフレーム化され、各フレームの音声データは、観測ベクトル a として、特徴抽出部2に順次供給される。特徴抽出部2では、フレーム化部1からの観測ベクトル a としての音声データが音響分析され、その特徴ベクトル y が求められる。さらに、特徴抽出部2では、求められた特徴ベクトル y に基づいて、特徴ベクトル空間における分布を表す特徴分布パラメータ Z が算出され、識別部3に供給される。識別部3では、特徴抽出部2からの特徴分布パラメータを用いて、所定数 K の単語それぞれに対応するクラスの識別関数の値が演算され、その関数値の最も大きいクラスが、音声の認識結果として出力される。

【0034】以上のように、観測ベクトル a としての音声データが、その特徴量の空間である特徴ベクトル空間における分布を表す特徴分布パラメータ Z に変換されるので、その特徴分布パラメータは、音声データに含まれるノイズの分布特性を考慮したものとなっており、さらに、そのような特徴分布パラメータに基づいて、音声 が認識されるので、認識率を、大きく向上させることが可能となる。

【0035】次に、図4は、図1の特徴抽出部2の構成例を示している。

【0036】観測ベクトル a は、パワースペクトラム分析器12に供給される。パワースペクトラム分析器12では、観測ベクトル a が、例えば、FFT（高速フーリエ変換）アルゴリズムによってフーリエ変換され、これにより、音声の特徴量であるパワースペクトラムが、特徴ベクトルとして求められる（抽出される）。なお、ここでは、1フレームの音声データとしての観測ベクトル a が、 D 個のコンポーネントからなる特徴ベクトル（ D 次元の特徴ベクトル）に変換されるものとする。

【0037】ここで、いま、第 t フレームの観測ベクトル $a(t)$ から得られる特徴ベクトルを $y(t)$ を表す。また、特徴ベクトル $y(t)$ のうち、真の音声のスペクトル成分を $x(t)$ と、ノイズのスペクトル成分を $u(t)$ と表す。この場合、真の音声のスペクトル成分 $x(t)$ は、次式で表される。

【0038】

【数1】

$$x(t) = y(t) - u(t)$$

・・・ (1)

但し、ここでは、ノイズが不規則な特性を有し、また、観測ベクトル $a(t)$ としての音声データが、真の音声成分に、ノイズを加算したものとなっているとしている。

【0039】ノイズ $u(t)$ の特性は不規則であるから、 $u(t)$ は確率変数であり、従って、式 (1) で表される $x(t)$ も確率変数となる。従って、ノイズのパワースペクトラムの確率密度関数が、例えば、図5

(A) に示すように表されるとき、式 (1) から、真の音声のパワースペクトラムの確率密度関数は、同図

(B) に示すように表される。即ち、真の音声のパワースペクトラムが、ある値である確率は、その値を、入力音声（入力信号）のパワースペクトラムから減算し、ノイズのパワースペクトラムが、その減算値となるときの確率に、真の音声の確率分布の面積が1になるような正規化係数（ファクタ）を乗じた値となる。なお、図5においては、 $u(t)$ 、 $x(t)$ 、 $y(t)$ のコンポーネント数が1 ($D=1$) としてある。

【0040】図4に戻り、パワースペクトラム分析器12で得られた特徴ベクトル $y(t)$ は、スイッチ13に供給される。スイッチ13は、音声区間検出部11の制御にしたがって、端子13aまたは13bのうちのいずれか一方を選択するようになされている。

【0041】即ち、音声区間検出部11では、音声区間（ユーザの発話がなされている期間）が検出される。ここで、音声区間の検出方法については、例えば、J.C. Junqua, B. Mark, and B. Reaves, "A Robust Algorithm for Word Boundary Detection in the Presence of Noise", IEEE Transaction Speech and Audio Processing, Vol. 2, No. 3, 1994などに、その詳細が開示されている。

【0042】なお、音声区間は、その他、例えば、音声認識装置に、所定のボタンを設けておき、発話を行っている間は、ユーザに、そのボタンを操作してもらうようにすることで認識するようすることも可能である。

【0043】音声区間検出部11は、音声区間においては、スイッチ13が端子13bを選択するように制御を行い、それ以外の区間（以下、適宜、非音声区間という）においては、スイッチ13が端子13aを選択するように制御を行う。

【0044】従って、非音声区間においては、スイッチ13は端子13bを選択し、これにより、パワースペクトラム分析器12の出力は、スイッチ13を介して、ノイズ特性算出器14に供給される。ノイズ特性算出器14では、非音声区間における、パワースペクトラム分析器12の出力から、音声区間におけるノイズの特性が求められる。

【0045】即ち、ここでは、ある音声区間におけるノイズのパワースペクトラム $u(t)$ の分布が、その音声区間の直前の非音声区間におけるノイズと同一であるとし、さらに、その分布が正規分布であるとして、ノイズ特性算出器14において、ノイズの平均値（平均ベクトル）と分散（分散マトリクス）が求められる。

【0046】具体的には、いま、音声区間の最初のフレームを第1フレーム ($t=1$) とすると、その200フレーム前から101フレーム前までの100フレーム分のパワースペクトラム分析器12の出力 $y(-200)$ 乃至 $y(-101)$ の平均ベクトル μ' と分散マトリクス Σ' とが、音声区間におけるノイズの特性として求め

10

20

30

40

50

られる。

【0047】ここで、平均ベクトル μ' と分散マトリクス Σ' は、次式にしたがって求めることができる。

【0048】

【数2】

$$\mu'(i) = \frac{1}{100} \sum_{t=-200}^{-101} y(t)(i)$$

$$\Sigma'(i,j) = \frac{1}{100} \sum_{t=-200}^{-101} (y(t)(i) - \mu'(i))(y(t)(j) - \mu'(j))$$

・・・(2)

但し、 $\mu'(i)$ は、平均ベクトル μ' の*i*番目のコンポーネントを表す($i=1, 2, \dots, D$)。また、 $y(t)(i)$ は、第*t*フレームの特徴ベクトルの*i*番目のコンポーネントを表す。さらに、 $\Sigma'(i, j)$ は、分散マトリクス Σ' の、第*i*行、第*j*列のコンポーネントを表す($j=1, 2, \dots, D$)。

【0049】ここで、計算量の低減のために、ノイズについては、特徴ベクトル y の各コンポーネントが、互いに無相関であると仮定する。この場合、次式に示すように、分散マトリクス Σ' は、対角成分以外は0となる。 20

【0050】

【数3】

$$\Sigma'(i,j) = 0, i \neq j$$

・・・(3)

【0051】ノイズ特性算出器14では、以上のようにして、ノイズの特性としての平均ベクトル μ' および平均値 Σ' が求められ、特徴分布パラメータ算出器15に供給される。

【0052】一方、音声区間においては、スイッチ13は端子13aを選択し、これにより、パワースペクトラム分析器12の出力、即ち、真の音声とノイズとを含む音声データとしての特徴ベクトル y は、スイッチ13を介して、特徴分布パラメータ算出器15に供給される。 30

特徴分布パラメータ算出器15では、パワースペクトラム分析器12からの特徴ベクトル y 、およびノイズ特性算出器14からのノイズの特性に基づいて、真の音声のパワースペクトラムの分布(推定値の分布)を表す特徴分布パラメータが算出される。

【0053】即ち、特徴分布パラメータ算出器15では、真の音声のパワースペクトラムの分布が正規分布であるとして、その平均ベクトル μ と分散マトリクス Ψ が、特徴分布パラメータとして、次式にしたがって計算される。

【0054】

【数4】

$$\begin{aligned} \mu(i) &= E[x(t)(i)] \\ &= E[y(t)(i) - u(t)(i)] \\ &= \int_0^{r(i)} (y(t)(i) - u(t)(i)) \frac{P(u(t)(i))}{\int_0^{r(i)} P(u(t)(i)) du(t)(i)} du(t)(i) \\ &= \frac{y(t)(i) \int_0^{r(i)} P(u(t)(i)) du(t)(i) - \int_0^{r(i)} u(t)(i) P(u(t)(i)) du(t)(i)}{\int_0^{r(i)} P(u(t)(i)) du(t)(i)} \\ &= y(t)(i) - \frac{\int_0^{r(i)} u(t)(i) P(u(t)(i)) du(t)(i)}{\int_0^{r(i)} P(u(t)(i)) du(t)(i)} \end{aligned}$$

・・・(4)

【数5】

$i=j$ のとき

$$\begin{aligned} \Psi(i,i) &= V[x(t)(i)] \\ &= E[(x(t)(i))^2] - (E[x(t)(i)])^2 \\ &= (E[(x(t)(i))^2] - (\mu(i))^2) \end{aligned}$$

$i \neq j$ のとき

$$\Psi(i,j) = 0$$

・・・(5)

【数6】

$$E \{ (x(t)(i))^2 \} = E \{ (y(t)(i) - u(t)(i))^2 \}$$

$$\begin{aligned} & - \int_0^{y(t)(i)} (y(t)(i) - u(t)(i))^2 \frac{P(u(t)(i))}{\int_0^{y(t)(i)} P(u(t)(i)) du(t)(i)} du(t)(i) \\ & - \frac{1}{\int_0^{y(t)(i)} P(u(t)(i)) du(t)(i)} \times \left\{ (y(t)(i))^2 \int_0^{y(t)(i)} P(u(t)(i)) du(t)(i) \right. \\ & \quad - 2y(t)(i) \int_0^{y(t)(i)} u(t)(i) P(u(t)(i)) du(t)(i) \\ & \quad \left. + \int_0^{y(t)(i)} (u(t)(i))^2 P(u(t)(i)) du(t)(i) \right\} \\ & - (y(t)(i))^2 - 2y(t)(i) \frac{\int_0^{y(t)(i)} u(t)(i) P(u(t)(i)) du(t)(i)}{\int_0^{y(t)(i)} P(u(t)(i)) du(t)(i)} \\ & \quad + \frac{\int_0^{y(t)(i)} (u(t)(i))^2 P(u(t)(i)) du(t)(i)}{\int_0^{y(t)(i)} P(u(t)(i)) du(t)(i)} \end{aligned}$$

・・・ (6)

【数 7】

$$P(u(t)(i)) = \frac{1}{\sqrt{2\pi} \sum (i,i)} e^{-\frac{1}{2\sum (i,i)} (u(t)(i) - \mu^*(i))^2}$$

・・・ (7)

【0055】ここで、 $\xi(t)(i)$ は、第 t フレームにおける平均ベクトル $\xi(t)$ の i 番目のコンポーネントを表す。また、 $E[]$ は、 $[]$ 内の平均値を意味する。 $x(t)(i)$ は、第 t フレームにおける真の音声の
30 パワースペクトラム $x(t)$ の i 番目のコンポーネントを表す。さらに、 $u(t)(i)$ は、第 t フレームにおけるノイズのパワースペクトラムの i 番目のコンポーネントを表し、 $P(u(t)(i))$ は、第 t フレームにおけるノイズのパワースペクトラムの i 番目のコンポーネントが $u(t)(i)$ である確率を表す。ここでは、ノイズの分布として正規分布を仮定しているので、 $P(u(t)(i))$ は、式 (7) に示したように表される。

【0056】また、 $\Psi(t)(i, j)$ は、第 t フレーム
40 における分散 $\Psi(t)$ の、第 i 行、第 j 列のコンポーネントを表す。さらに、 $V[]$ は、 $[]$ 内の分散を表す。

【0057】特徴分布パラメータ算出器 15 では、以上のようにして、各フレームごとに、平均ベクトル ξ および分散マトリクス Ψ が、真の音声の特徴ベクトル空間上での分布（ここでは、真の音声の特徴ベクトル空間上での分布が正規分布であると仮定した場合の、その分布）を表す特徴分布パラメータとして求められる。

【0058】その後、音声区間が終了すると、スイッチ 50

13 は端子 13b を選択し、また、特徴分布パラメータ算出器 15 は、音声区間の各フレームにおいて求めた特徴分布パラメータを、識別部 3 に出力する。即ち、いま、音声区間が T フレームであったとし、その T フレームそれぞれにおいて求められた特徴分布パラメータを、 $z(t) = \{\xi(t), \Psi(t)\}$ と表すと ($t=1, 2, \dots, T$)、特徴分布パラメータ算出器 15 は、特徴分布パラメータ (系列) $Z = \{z(1), z(2), \dots, z(T)\}$ を、識別部 3 に供給する。

【0059】そして、特徴抽出部 2 では、以下、同様の処理が繰り返される。

【0060】次に、図 6 は、図 1 の識別部 3 の構成例を示している。

【0061】特徴抽出部 2 (特徴分布パラメータ算出器 15) からの特徴分布パラメータ Z は、 K 個の識別関数演算部 21、乃至 21_K に供給される。識別関数演算部 21_k は、 K 個のクラスの中の k 番目に対応する単語を識別するための識別関数 $g_k(Z)$ を記憶しており ($k=1, 2, \dots, K$)、特徴抽出部 2 からの特徴分布パラメータ Z を引数として、識別関数 $g_k(Z)$ を演算する。

【0062】ここで、識別部 3 では、例えば、HMM (Hidden Markov Model) 法を用いて、クラスとしての単語の識別 (認識) が行われるようになされている。

【0063】なお、ここでは、例えば、図 7 に示すような HMM を用いる。即ち、ここでは、HMM は、 H 個の状態 q_1 、乃至 q_H を有しており、状態の遷移は、自身への遷移と、右隣の状態への遷移のみが許されている。また、初期状態は、最も左の状態 q_1 とされ、最終状態は、最も右の状態 q_H とされており、最終状態 q_H からの

状態遷移は禁止されている。このように、自身よりも左にある状態への遷移のないモデルは、left-to-rightモデルと呼ばれ、音声認識では、一般に、left-to-rightモデルが用いられる。

【0064】いま、HMMのkクラスを識別するためのモデルを、kクラスモデルというすると、kクラスモデルは、例えば、最初に状態 q_1 にいる確率（初期状態確率） $\pi_k(q_1)$ 、ある時刻（フレーム）tにおいて、状態 q_1 にいて、次の時刻t+1において、状態 q_1 に状態遷移する確率（遷移確率） $a_k(q_1, q_1)$ 、および状態 q_1 から状態遷移が生じるときに、その状態 q_1 が、特徴ベクトルOを出力する確率（出力確率） $b_k(q_1)$ （O）によって規定される（ $h=1, 2, \dots, H$ ）。

【0065】そして、ある特徴ベクトル系列 $O_1, O_2,$

$$g_k(Z) = \max_{q_1, q_2, \dots, q_T} \pi_k(q_1) \cdot b_k(q_1)(z_1) \cdot a_k(q_1, q_2) \cdot b_k(q_2)(z_2)$$

$$\dots a_k(q_{T-1}, q_T) \cdot b_k(q_T)(z_T)$$

・・・(8)

【0068】ここで、 $b_k(q_1)(z_1)$ は、出力が z_1 で表される分布であるときの出力確率を表す。状態遷移時に各特徴ベクトルを出力する確率である出力確率 $b_k(s)(O_1)$ には、ここでは、例えば、特徴ベクトル空間上のコンポーネントどうしの間に相関がないものとして、正規分布関数が用いられている。この場合、入力 z_1 で表される分布であるとき、出力確率 $b_k(q_1)(z_1)$

$$b_k(s)(z_1) = \int P^1(t)(x) P_k^m(s)(x) dx$$

$$= \prod_{i=1}^D P(s)(i)(\xi(t)(i), \Psi(t)(i, i))$$

$$k=1, 2, \dots, K; s=q_1, q_2, \dots, q_T; T=1, 2, \dots, T$$

・・・(9)

但し、式(9)における積分の積分区間は、D次元の特徴ベクトル空間（ここでは、パワースペクトラム空間）の全体である。

【0070】また、式(9)において、 $P(s)(i)$

$$P(s)(i)(\xi(t)(i), \Psi(t)(i, i))$$

$$= \frac{1}{\sqrt{2\pi(\Sigma_k(s)(i, i) + \Psi(t)(i, i))}} e^{-\frac{(\mu_k(s)(i) - \xi(t)(i))^2}{2(\Sigma_k(s)(i, i) + \Psi(t)(i, i))}}$$

・・・(10)

但し、 $\mu_k(s)(i)$ は、平均ベクトル $\mu_k(s)$ のi番目のコンポーネントを、 $\Sigma_k(s)(i, i)$ は、分散マトリクス $\Sigma_k(s)$ の、第i行第i列のコンポーネントを、それぞれ表す。そして、kクラスモデルの出力確率は、これらによって規定される。

【0072】なお、HMMは、上述したように、初期状態確率 $\pi_k(q_1)$ 、遷移確率 $a_k(q_1, q_1)$ 、および出力確率 $b_k(q_1)(O)$ によって規定されるが、これ

・・・が与えられた場合に、例えば、そのような特徴ベクトル系列が観測される確率（観測確率）が最も高いモデルのクラスが、その特徴ベクトル系列の認識結果とされる。

【0066】ここでは、この観測確率が、識別関数 $g_k(Z)$ によって求められるようになされている。即ち、ここでは、識別関数 $g_k(Z)$ は、特徴分布パラメータ（系列） $Z = \{z_1, z_2, \dots, z_T\}$ に対する最適状態系列（最適な状態の遷移のしていき方）において、そのような特徴分布パラメータ（系列） $Z = \{z_1, z_2, \dots, z_T\}$ が観測される確率を求めるものとして、次式で与えられる。

【0067】

【数8】

(s) (z_1) は、平均ベクトル $\mu_k(s)$ と分散マトリクス $\Sigma_k(s)$ とによって規定される確率密度関数 $P_k^m(s)(x)$ 、および第tフレームの特徴ベクトル（ここでは、パワースペクトラム）xの分布を表す確率密度関数 $P^1(t)(x)$ を用いて、次式により求めることができる。

【0069】

【数9】

($\xi(t)(i), \Psi(t)(i, i)$)は、次式で表される。

【0071】

【数10】

らは、学習用の音声データから特徴ベクトルを算出し、その特徴ベクトルを用いて、あらかじめ求めておくようにする。

【0073】ここで、HMMとして、図7に示したものをを用いる場合には、常に、最も左の状態 q_1 から遷移が始まるので、初期状態確率は、状態 q_1 についてのもののみ1とされ、他の状態についてのものはすべて0とされる。また、出力確率は、式(9)および(10)から明らかなように、 $\Psi(t)(i, i)$ を0とすると、特

微ベクトルの分散を考慮しない場合の連続HMMにおける出力確率に一致する。

【0074】なお、HMMの学習方法としては、例えば、Baum-Welchの再推定法などが知られている。

【0075】図6の識別関数演算部21は、kクラスモデルについて、あらかじめ学習により求められた初期状態確率 $\pi_k(q_1)$ 、遷移確率 $a_k(q_i, q_j)$ 、および出力確率 $b_k(q_i)$ (O)によって規定される式

(8)の識別関数 $g_k(Z)$ を記憶しており、特徴抽出部2からの特徴分布パラメータ Z を引数として、識別関数 $g_k(Z)$ を演算し、その関数値(上述した観測確率) $g_k(Z)$ を、決定部22に出力する。

【0076】決定部22では、識別関数演算部21乃至21、それぞれからの関数値 $g_k(Z)$ に対して、例えば、次式に示す決定規則を用いて、特徴分布パラメータ Z 、即ち、入力された音声に属するクラスが識別(決定)される(入力された音声がいずれかのクラスに分類される)。

【0077】

【数11】

$$C(Z) = C_k, \text{ if } g_k(Z) = \max_i \{g_i(Z)\}$$

・・・(11)

但し、 $C(Z)$ は、特徴分布パラメータ Z が属するクラスを識別する識別操作(処理)を行う関数を表す。また、式(11)の第2式の右辺における \max は、それに続く関数値 $g_i(Z)$ (但し、ここでは、 $i=1, 2, \dots, K$)の最大値を表す。

【0078】決定部22は、式(11)にしたがって、クラスを決定すると、それを、入力された音声の認識結果として出力する。

【0079】次に、図8は、図1の特徴抽出部2の他の構成例を示している。なお、図中、図4における場合と対応する部分については、同一の符号を付してある。即ち、この特徴抽出部2は、ノイズ特性算出器14または特徴分布パラメータ算出器15に代えて、ノイズバッファ31または特徴分布パラメータ算出器32がそれぞれ設けられている他は、基本的に、図4における場合と同様に構成されている。

【0080】この場合、ノイズバッファ31には、例えば、非音声区間における、パワースペクトラム分析器12の出力が、ノイズのパワースペクトラムとして一時記憶される。即ち、例えば、音声区間の200フレーム前から101フレーム前までの100フレーム分のパワースペクトラム分析器12の出力 $y(-200)$ 乃至 $y(-101)$ を、 $w(1)=y(-200)$ 、 $w(2)=y(-199)$ 、 \dots 、 $w(100)=y(-101)$ とそれぞれすると、ノイズバッファ31では、 $w(1)$ 乃至 $w(100)$ がノイズのパワースペクトラムとして記憶される。

【0081】そして、この100フレーム分のノイズのパワースペクトラム $w(n)$ は($w=1, 2, \dots, N$; 但し、ここでは、 $N=100$)、音声区間となると、特徴分布パラメータ算出器32に出力される。

【0082】特徴分布パラメータ算出器32は、音声区間となり、ノイズバッファ31からノイズのパワースペクトラム $w(n)$ ($w=1, 2, \dots, N$)を受信すると、真の音声のパワースペクトラムの分布(真の音声のパワースペクトラムの推定値の分布)を、例えば正規分布と仮定して、その分布を規定する平均ベクトル $\xi(t)$ と分散マトリクス $\Sigma(t)$ を、例えば、次式にしたがって算出する。

【0083】

【数12】

$$\xi(t)(i) = E\{x(t)(i)\}$$

$$= \frac{1}{N} \sum_{n=1}^N \{y(t)(i) - w(n)(i)\}$$

$$\Psi(t)(i,j) = \frac{1}{N} \sum_{n=1}^N \{ \{y(t)(i) - w(n)(i) - \xi(t)(i)\} \times \{y(t)(j) - w(n)(j) - \xi(t)(j)\} \}$$

$$j=1,2,\dots,D : j=1,2,\dots,D$$

・・・(12)

但し、 $w(n)(i)$ は($w(n)(j)$ も同様)、 n 番目のノイズのパワースペクトラム $w(n)$ の i 番目のコンポーネントを表す。

【0084】特徴分布パラメータ算出器32は、以上のようにして、平均ベクトル $\xi(t)$ と分散マトリクス $\Sigma(t)$ を、フレームごとに求め、これらの組合せを、特徴分布パラメータ z_t とし、音声区間における特徴分布パラメータ $Z=\{z_1, z_2, \dots, z_t\}$ を、識別部3に出力する。

【0085】図4においては、ノイズのパワースペクトラムのコンポーネントどうしが無相関であることを仮定したが、図8では、そのような仮定をせずに、特徴分布パラメータが求められるので、より精度の良い特徴分布パラメータを得ることができる。

【0086】次に、以上においては、特徴ベクトル(特徴量)として、パワースペクトラムを用いるようにしたが、特徴ベクトルとしては、その他、例えば、ケプストラムなどを用いることも可能である。

【0087】即ち、いま、あるフレーム t の真の音声のケプストラムを $x^*(t)$ と表し、その分布(ケプストラムの推定値の分布)を、例えば、正規分布と仮定すると、第 t フレームの特徴ベクトル(ここでは、ケプストラム) x^* の分布を表す確率密度関数 $P^*(t)(x^*)$ を規定する平均ベクトル $\xi^*(t)$ と分散マトリクス $\Psi^*(t)$ は、次式にしたがって求めることができる。

【0088】

【数13】

$$\xi^c(t)(i) = \frac{1}{N} \sum_{n=1}^N x^c(t)(n)(i)$$

$$i=1,2,\dots,D$$

$$\Psi^c(t)(i,j) = \frac{1}{N} \sum_{n=1}^N (x^c(t)(n)(i) - \xi^c(t)(i))(x^c(t)(n)(j) - \xi^c(t)(j))$$

$$i=1,2,\dots,D : j=1,2,\dots,D$$

・・・ (13)

但し、 $\xi^c(t)(i)$ は、平均ベクトル $\xi^c(t)$ の i 番目のコンポーネントを表し、 $\Psi^c(t)(i,j)$ は、分散マトリクス $\Psi^c(t)$ の、第 i 行、第 j 列のコンポーネントを表す。また、 $x^c(t)(n)(i)$

は、次式で表されるケブストラム $x^c(t)(n)$ の i 番目のコンポーネントを表す。

【0089】

【数14】

$$x^c(t)(n) = Cx^L(t)(n)$$

$$x^L(t)(n) = (x^L(t)(n)(1), x^L(t)(n)(2), \dots, x^L(t)(n)(D))$$

$$x^L(t)(n)(i) = \log(y(t)(i) - w(n)(i)) \quad i=1,2,\dots,D$$

・・・ (14)

但し、式 (14) の第 1 式における C は、離散コサイン変換を行う DCT (Discrete Cosine Transform) マトリクスである。

【0090】特徴ベクトルとしてケブストラムを用いる場合には、図 1 の特徴抽出部 2 において、以上のようにして、平均ベクトル $\xi^c(t)$ と分散マトリクス $\Psi^c(t)$ を、フレームごとに求め、これらの組合せ $\{\xi^c(t), \Psi^c(t)\}$ を、特徴分布パラメータ z_i^c とし、音声区間における特徴分布パラメータ $Z^c = \{z_1^c, z_2^c, \dots, z_T^c\}$ を、識別部 3 に出力する

ようにすれば良い。

【0091】この場合、識別部 3 において識別関数 $g_i(Z^c)$ を計算するのに用いる出力確率 $b_i^c(s)(z_i^c)$ は、ケブストラム空間上の分布を表す確率として、平均ベクトル $\mu_{i^c}(s)$ と分散マトリクス $\Sigma_{i^c}(s)$ とによって規定される確率密度関数 $P_{i^c}(s)(x^c)$ 、および第 t フレームの特徴ベクトル (ここでは、ケブストラム) x^c の分布を表す確率密度関数 $P^c(t)(x^c)$ を用いて、次式により求めることができる。

【0092】

【数15】

$$b_k^c(s)(z_i^c) = \int P^c(x^c) P_k^m(s)(x^c) dx^c$$

$$= \frac{e^{-\frac{1}{2}(\xi^c(t) - \mu_{i^c}(s))^T (\Psi^c(t) + \Sigma_{i^c}(s))^{-1} (\xi^c(t) - \mu_{i^c}(s))}}{(2\pi)^{\frac{D}{2}} |\Psi^c(t) + \Sigma_{i^c}(s)|^{\frac{1}{2}}}$$

・・・ (15)

但し、式 (15) における積分の積分区間は、 D 次元の特徴ベクトル空間 (ここでは、ケブストラム空間) の全体である。また、 $(\xi^c(t) - \mu_{i^c}(s))^T$ は、ベクトル $(\xi^c(t) - \mu_{i^c}(s))$ の転置を表す。

【0093】以上のように、ノイズの分布特性を考慮した特徴分布パラメータを求め、その特徴分布パラメータを用いて音声認識を行うようにしたので、認識率を向上させることができる。

【0094】ここで、音声の特徴量としてケブストラムを、識別部 3 における音声認識アルゴリズムとして HMM 法を、それぞれ用いて、特徴分布パラメータによる音声認識 (単語認識) 実験を行った場合と、スペクトラムサブトラクションによる音声認識実験を行った場合の認識率を、表 1 に示す。

【0095】

【表1】

音声入力環境	認識率 (%)	
	SS法	本発明
アイドリング+背景音楽	72	86
市街地走行	85	90
高速道路走行	57	69

但し、認識対象単語は 5000 単語で、話者は不特定話者とした。また、発話は、自動車がアイドリング状態にあり、かつ BGM (Back Ground Music) が流れている環境、自動車でも市街地を走行している環境、および自動車でも高速道路を走行している環境の 3 つの環境で行った。

【0096】表 1 から明らかなように、いずれの環境下においても、特徴分布パラメータによる音声認識結果の方が、認識率が高くなっている。

【0097】以上、本発明を適用した音声認識装置について説明したが、このような音声認識装置は、例えば、

音声入力可能なカーナビゲーション装置その他各種の装置に適用可能である。

【0098】なお、本実施の形態では、ノイズの分布特性を考慮した特徴分布パラメータを求めるようにしたが、このノイズには、例えば、発話を行う環境下における外部からのノイズの他、例えば、電話回線その他の通信回線を介して送信されてくる音声の認識を行う場合には、その通信回線の特性なども含まれる。

【0099】また、本発明は、例えば、特定話者認識を行う場合において、その特定話者の学習を行うときなどにも適用可能である。この場合、学習速度を向上させることが可能となる。

【0100】さらに、本発明は、音声認識の他、画像認識その他のパターン認識を行う場合にも適用可能である。例えば、画像認識においては、画像を撮影するカメラのレンズの特性や、気象状況などをノイズとして、その分布特性を考慮した特徴分布パラメータを用いるようにすることで、画像の認識率を向上させることが可能となる。

【0101】また、本実施の形態では、パワースペクトラム空間またはケプストラム空間上の分布を表す特徴分布パラメータを求めるようにしたが、分布を求める空間としては、その他、線形予測係数や、隣接フレームのケプストラムの差分、ゼロクロスなどの空間を用いることも可能である。

【0102】さらに、本実施の形態では、音声の1つ(1種類)の特徴量の空間における分布を表す特徴分布パラメータを求めるようにしたが、その他、音声の複数の特徴量の空間について、それぞれ、特徴分布パラメータを求めるようにすることも可能である。あるいは、音声の複数の特徴量の空間のうちの1以上について特徴分布パラメータを求め、その特徴分布パラメータと、残りの特徴量の空間における特徴ベクトルとの両方を用いて、音声認識を行うようにすることも可能である。

【0103】また、本実施の形態では、特徴ベクトル空間上における特徴ベクトル(真の音声の特徴ベクトルの推定値)の分布が正規分布であるとして、そのような分布を表す特徴分布パラメータを用いるようにしたが、特徴分布パラメータとしては、その他、例えば、対数正規確率分布や、離散確率分布、ファジー分布などを表すものを用いることも可能である。

【0104】さらに、本実施の形態では、識別部3において、出力確率が正規分布で表されるHMMを用いて、クラスの識別を行うようにしたが、識別部3には、その他、例えば、出力確率が混合正規確率分布や離散分布などで表されるHMMや、正規確率分布関数、対数確率分布関数、多項式関数、ニューラルネットワークなどを用

いて、クラスの識別を行わせるようにすることも可能である。

【0105】

【発明の効果】請求項1に記載の特徴抽出装置および請求項7に記載の特徴抽出方法によれば、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータが算出される。従って、例えば、入力データにノイズが含まれる場合に、そのノイズの分布特性を考慮したパラメータを得ることができ

る。

【0106】請求項8に記載のパターン認識装置および請求項24に記載のパターン認識方法によれば、入力データの特徴量の空間に、入力データを写像したときに得られる分布を表す特徴分布パラメータが算出され、その特徴分布パラメータが、所定数のクラスのうちのいずれかに分類される。例えば、入力データにノイズが含まれる場合に、そのノイズの分布特性を考慮したパラメータを得ることができるので、入力データの認識率を向上させることが可能となる。

【図面の簡単な説明】

【図1】本発明を適用した音声認識装置の一実施の形態の構成例を示すブロック図である。

【図2】図1のフレーム化部1の処理を説明するための図である。

【図3】図1の特徴抽出部2の処理を説明するための図である。

【図4】図1の特徴抽出部2の構成例を示すブロック図である。

【図5】ノイズおよび真の音声のパワーの確率密度関数を示す図である。

【図6】図1の識別部3の構成例を示すブロック図である。

【図7】HMMを示す図である。

【図8】図1の特徴抽出部2の他の構成例を示すブロック図である。

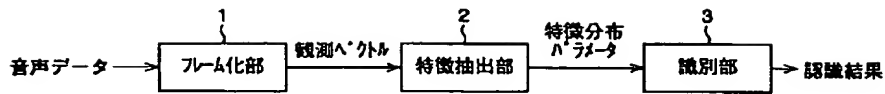
【図9】従来のパターン認識装置の一例の構成を示すブロック図である。

【図10】図9の特徴抽出部101の処理を説明するための図である。

【符号の説明】

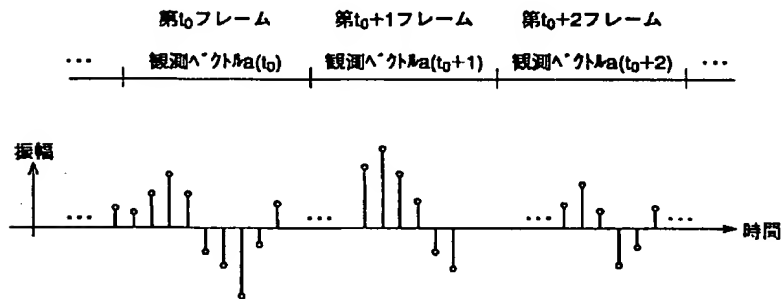
1 フレーム化部, 2 特徴抽出部(算出手段),
3 識別部(分類手段), 11 音声区間検出部,
12 パワースペクトラム分析器, 13 スイッチ,
13 a, 13 b 端子, 14 ノイズ特性算出器,
15 特徴分布パラメータ算出器, 21,乃至21,
識別関数演算部, 22 決定部, 31 ノイズバッ
ファ, 32 特徴分布パラメータ算出器

【図 1】

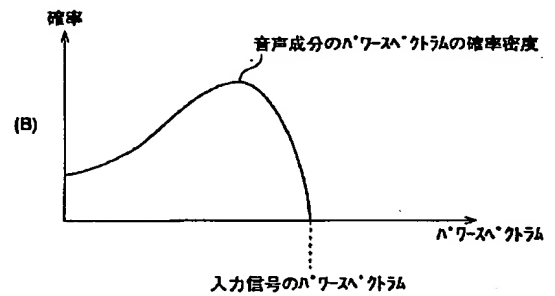
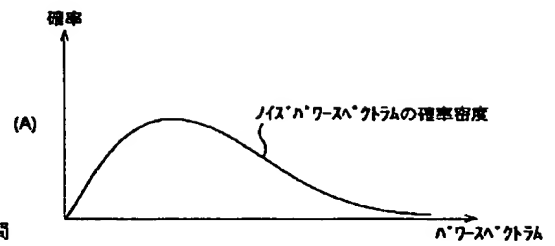


音声認識装置

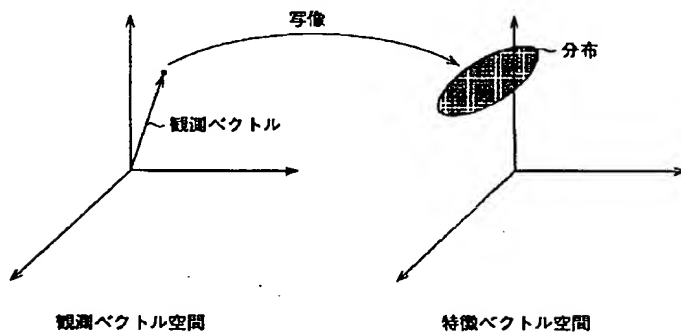
【図 2】



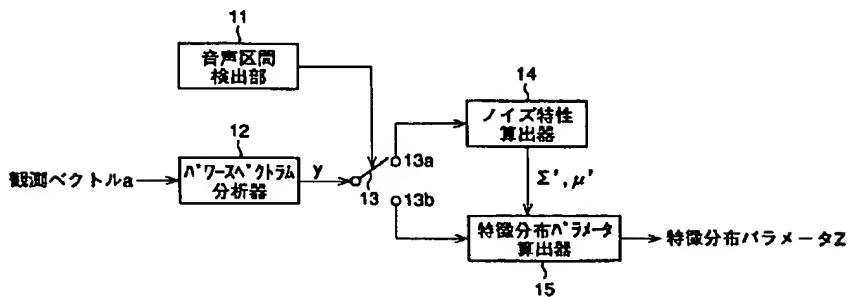
【図 5】



【図 3】

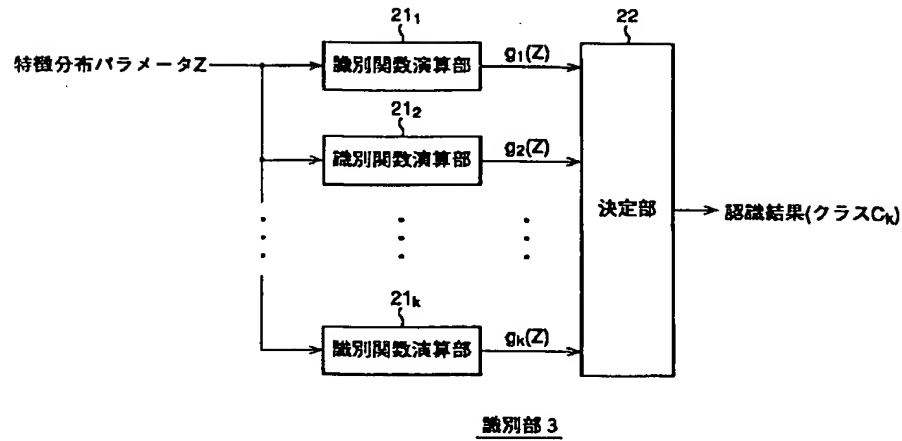


【図 4】

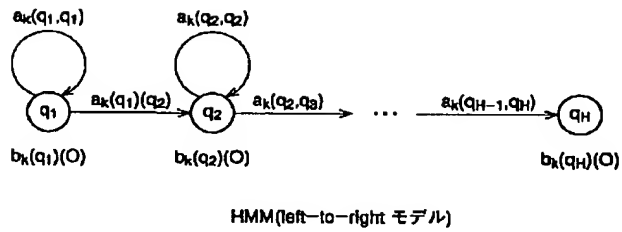


特徴抽出部 2

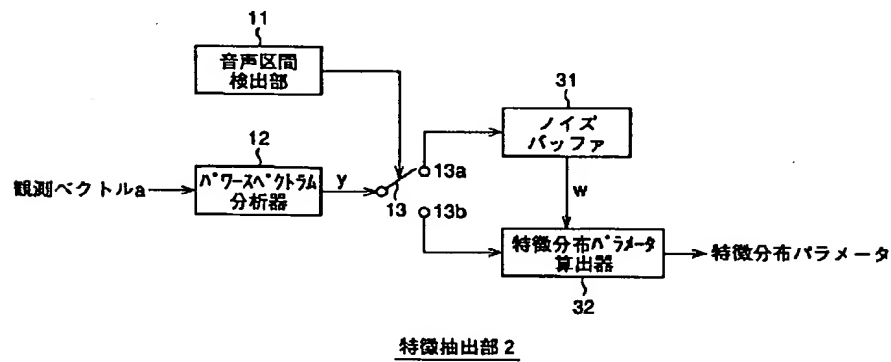
【図 6】



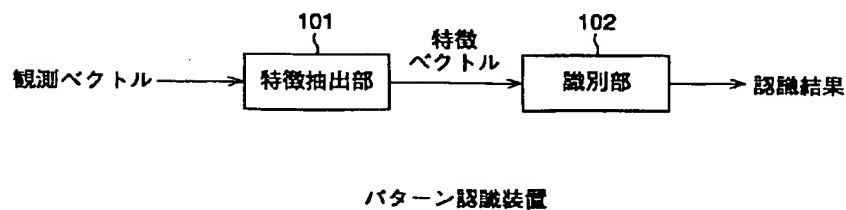
【図 7】



【図 8】



【図 9】



【図 1 0】

